



Progressing Towards Open Textbooks Learning Analytics System

Deepak Prasad^{a,b}, Rajneel Totaram^b, Tsuyoshi Usagawa^a

^aKumamoto University, Japan

^bThe University of the South Pacific, Fiji

ABSTRACT

Textbook prices have been soaring at an unprecedented pace for the last four decades with no signs that this trend will end anytime soon. Several studies have suggested that a solution to this problem comes in the form of open textbooks. As a result, the growth of open textbooks is rapid and sustained. However, though the advent of open textbooks is encouraging, whether, how, and to what extent students are using their open textbooks remains unclear. Learning analytics for open textbooks can provide answers to these questions plus many others, and thereby offers the potential for improving planning, development, monitoring, evaluation and revision of open textbooks.

Learning analytics applied to open textbooks has received little attention to date. This on the horizon paper presents and describes developmental work of a method to collect data produced as a result of students' online and offline interactions with their open textbooks, the first part of a three-step process of learning analytics (the remaining two being data processing and reporting functionalities). The paper concludes with a presentation of future work, in line with the nature of this paper, which is work-in-progress towards developing learning analytics system for open textbooks.

Keywords: open textbooks; learning analytics; EPUB reader; online and offline data recording; data synchronisation.

Introduction

The global context of higher education is undergoing monumental transformation impelled by a vast array of issues, with much concern connected with rising cost – from tuition and fees to living expenses, data plans and textbook costs – which is often seen as a major hindrance to acquiring higher education credentials (Driscoll, Comm, & Mathaisel, 2013; Yang & McCall, 2014). A major factor of these financial issues, and one that stands out due to the press it receives globally, is the rapid increase in the price of publisher textbooks, which are often a required component of higher education courses (Prasad & Usagawa, 2014) and thus “represent a significant portion of the overall cost of higher education” (Hilton, Robinson, Wiley, & Ackerman, 2014, p. 67).

The price of publishers' textbooks has risen a staggering 812% since 1978 (Perry, 2012). Impacts of this price increase have been reported to be quite dramatic, with studies revealing that two-thirds to three-quarters of college students find affording textbooks difficult (N. Allen, 2011; Senack, 2014). Several studies have also found that the degree of unaffordability is reaching the point that it impairs learning (Acker, 2011; N. Allen, 2011; Graydon, Urbach-Buholz, & Kohen, 2011; Morris-Babb & Henderson, 2012; Senack, 2014). Despite these problems, at present there is no indication of any slowdown in price increases.

Open textbooks, a form of open educational resource, hold potential to mitigate the crisis in textbook affordability. While open textbooks are similar to traditional publishers' textbooks in terms of content, they are generally available free of charge both in a variety of digital formats and at low cost for print copies (Hilton, Gaudet, Clark, Robinson, & Wiley, 2013). In fact, several recent studies have shown that open textbooks can achieve significant cost savings with no negative effect on learning (G. Allen, Guzman-Alvarez, Molinaro, & Larsen, 2015; Robinson, Fischer, Wiley, & Hilton, 2014; Senack, 2014; Wiley, Hilton, Ellington, & Hall, 2012). For instance, specifically in monetary terms, survey evidence presented in Senack (2014) shows that open textbooks could save college students an average of \$100 per course. Findings such as these have led to a growing acceptance of open textbooks within higher educational institutions as an affordable alternative to expensive publishers' textbooks.

The rapid growth in the popularity of open textbooks worldwide is accompanied by equally rapid growth in production. Innovation results in unfamiliarity, however, and surprisingly little data exists regarding how exactly students use these textbooks or even whether they use them at all. Answers to these questions and others can be obtained by applying learning analytics to open textbooks. Fortunately, open textbooks, by their very nature of being digital, offer opportunities to enable learning analytic techniques to track students' interactions with their open textbooks, thereby providing useful new insights for all involved in open textbooks to enable improvement in their planning, development, monitoring, evaluation and revision.

Progressing Towards Open Textbooks Learning Analytics System

Little work has been done so far towards the development of a learning analytics system for open textbooks despite its obvious usefulness. To this end, this paper presents developmental work of a method for recording and synchronizing data – the first part of learning analytics 3-step process: data collection, data analysis and data reporting – produced as a result of students' online and offline interactions with their open textbooks. The developed method is designed to support open textbooks in the EPUB (short for electronic publication) format, a format that has become the international standard for digital textbooks. The prototype was tested successfully with a group of university students. While the work presented in this paper represents initial work towards an open textbook analytic system (planned future steps include data processing and reporting functionalities), it is still quite useful in its current form as the techniques described can be applied to capture data which can then be used for analysis purposes with a program of choice, and represents a major step in recording and synchronising data generated from both online and offline interactions with digital books. The following section outlines a brief review of literature on learning analytics.

An overview of learning analytics

The concept of learning analytics has been defined in several different ways; an oft-cited definition is “the measurement, collection, analysis and reporting of data about learners and their contexts, for the purposes of understanding and optimizing learning and the environments in which it occurs” (Siemens & Long, 2011, p. 34). Experts assert that learning analytics is primarily concerned with learning (Clow, 2013; Gašević, Dawson, & Siemens, 2015) and is “poised to benefit students in previously impossible ways” (Willis, 2014). It has also been pointed out that learning analytics provides the opportunity to “ask new questions and to ask old questions in new ways” (Haythornthwaite, Laat, & Dawson, 2013, p. 1374) to generate both information (What happened? What is happening now? What will happen?) and insight (How and why did it happen? What's the next best action? What's the best/worst that can happen?) (Davenport, Harris, & Morison, 2010).

Several uses of learning analytics have been identified and applied over the years, and the list continues to grow. Some of the recently reported uses include early-warning systems to identify struggling students in need of additional support; assessing the quality of online postings; visualising learner interactions; sending automated feedback messages; recommender systems for learning; intelligent tutoring systems; and identifying effective course designs (Arnold & Pistilli, 2012; Brooks, Greer, & Gutwin, 2014; Ferguson & Shum, 2011; Fritz, 2013; Gómez-Aguilar, Hernández-García, García-Peñalvo, & Therón, 2015; Jayaprakash, Moody, Lauría, Regan, & Baron, 2014; Manouselis, Drachler, Vuorikari, Hummel, & Koper, 2011; Scheffel et al., 2011; Tanes, Arnold, King, & Remnet, 2011; Wise, Zhao, & Hausknecht, 2014). With such a variety of applications, implementation of learning analytics also comes with “a number of legal, risk and ethical issues that should be taken into account” (Scheffel, Drachler, Stoyanov, & Specht, 2014, p. 128), particularly the need to maintain privacy of student data. In regards to these concerns, Pardo and Siemens (2014) have identified four principles: (1) transparency, (2) student control over data, (3) security, and (4) accountability and assessment, to inform decisions about how to comply with privacy-related matters.

From a technical perspective, learning analytics comprises three essential stages: (1) data collection, (2) data analysis, and (3) data reporting. A description of each stage synthesised from Brown (2011, p. 1) is presented below.

- Data collection: The first step to unlocking the potential of learning analytics. This involves using programs, scripts, and other methods to gather data from a single or multiple sources, and the data can be structured or unstructured. The specific design of the collection activity is informed by the goals of the learning analytics project.
- Data analysis: Unstructured data is usually given some kind of structure prior to analysis. The data is subjected to an appropriate combination of qualitative and quantitative analysis.
- Data reporting: The results of the analysis are reported using a combination of visualisations, tables, charts, and other kinds of information display.

Development of a prototype of a data collection system

This section describes the methods and techniques that were applied in the development of a data recording prototype as an initial step towards the development of a learning analytics system for open textbooks in electronic publication (EPUB) format; a free and open e-book standard by the International Digital Publishing (for more on EPUB, see <http://idpf.org/epub>). An EPUB file format is basically a zip file that contains a defined set on contents with the .epub extension that can be read using any EPUB reader application.

EPUB reader retrofitting strategies

Considering that EPUB reader applications are essential for viewing files in EPUB format, our underlying idea was to create a method for using an EPUB reader to facilitate the data collection process, particularly in terms of recording user interaction data, and then sending that data to the central database for storage and processing. To this end, it was necessary that all users use the same EPUB reader application to view the textbook; to ensure this, the EPUB file of the textbook was embedded into the EPUB reader application and then distributed as a single package.

EPUB.js reader selected for customisation

Rather than building from scratch, an existing open source EPUB reader for the web, EPUB.js (<https://github.com/futurepress/epub.js>), was chosen to be customised particularly because its web-based nature enabled use of appropriate web technologies for local data storage and transfer of these data to the central database.

Recording data

The EPUB.js in its original form already included capabilities to record user's click actions and their annotation notes within the local storage of the user's web browser running the EPUB.js reader application. To gather additional data, this functionality was extended to record other user data, such as browser used, type of device used, and IP address. However, user-side data storage was practically unusable for analysis purposes unless synced to a database for further processing, and the EPUB.js did not include such a function. This necessitated the customisation described below.

Synchronising data

In order to sync data stored within a browser's local storage to the central database server, a network-sensing feature was incorporated within the EPUB.js. The purpose of this feature was to make regular checks for connectivity to the central database server. The network detection was executed every 60 seconds. As soon as the connection was established, the data from the local storage was sent to the central database server and the local storage was cleared. This element is key to enable the transfer of offline interactional data to the central database server.

Tracking individual users

To uniquely track users, a simple authentication system was integrated in the EPUB.js reader. For this, when first using the modified EPUB.js reader application, users were prompted to enter their name and user ID. This information was then stored locally in the browser and all subsequent interactional data sent to the central database server were tagged under these individual user details.

Preparing books for online/offline access

The EPUB file of the open textbook and the customised EPUB.js reader were bundled together as a single package (hereinafter cited as hybrid-EPUB reader app), to be delivered in both online and offline modes.

For online access, the hybrid-EPUB reader app was hosted on a web server which users accessed using an URL. This enabled all user interaction to directly sync to the central database due to internet connectivity.

For offline consumption, the hybrid-EPUB reader app had to be saved on the user's local machine and accessed locally. To allow users to correctly save the hybrid-EPUB reader app onto their computers and access the correct file, an application installer was created for Windows-based users. The installer would save the application files in the correct location and add a shortcut on the computer. This allowed users to easily access the book. However, offline usage was restricted to Mozilla Firefox. This browser was chosen because the hybrid-EPUB reader app uses Javascript to send locally stored data to the central database, and most web browsers other than Firefox, considering it a security risk, block client side Javascript and cross-site scripting. Further investigation in this area is required.

Central database for data storage

The data storage server was a combination of a simple PHP script and a MySQL database. The MySQL database was used to record user interaction data. A PHP script waited for the data to be sent by the EPUB reader application. Data sent would immediately be received by the PHP script, which would conduct simple data validation and store them into the database.

Conclusion and future work

The work described above is only the first step of developing a learning analytics system for open textbooks. The work presented provides a foundation on which further research can take place. Our future work will involve extending this current work to include two additional steps: data processing and reporting functionalities. For data processing, analysis of both individual and aggregate data will be performed using SQL queries for various benchmarks such as total views per chapter, number of students versus chapters viewed, total bookmarks per chapter, user annotations (list of all notes) made per chapter, weekly user interaction, and online versus offline usage. To return processed information in a graphical and visual form, a dashboard format-reporting interface will be created using PHP and Javascript charting library. These extensions will be reported in detail in future publications.

Biographies

Deepak Prasad is a learning designer at the University of the South Pacific and a PhD student in the Graduate School of Science and Technology, Kumamoto University. His research is on open textbooks: teacher acceptance, student preference, and learning analytics. He tweets as @deepakvprasad.

Rajneel Totaram is a learning systems developer at the University of the South Pacific, and is interested in the association between learning and technology.

Tsuyoshi Usagawa is a professor in e-learning at Kumamoto University, and is interested in e-learning contents and systems. Address for correspondence: Deepak Prasad, Centre for Flexible Learning, University of the South Pacific, Private Mail Bag, Suva, Fiji. Email: deepak.v.prasad@gmail.com

References

- Acker, S. R. (2011). Digital textbooks: A state-level perspective on affordability and improved learning outcomes. *Library Technology Reports*, 47(8), 41–52.
- Allen, G., Guzman-Alvarez, A., Molinaro, M., & Larsen, D. (2015). *Assessing the impact and efficacy of the open-access chemwiki textbook project*. Retrieved from <https://net.educause.edu/ir/library/pdf/elib1501.pdf>
- Allen, N. (2011). *High prices prevent college students from buying assigned textbooks*. Student PIRGs. Retrieved from <http://www.studentpirgs.org/news/ap/high-prices-prevent-college-students-buying-assigned-textbooks>
- Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: using learning analytics to increase student success. In S. B. Shum, D. Gašević, & R. Ferguson (Eds.), *2nd International Conference on Learning Analytics and Knowledge (LAK '12)* (pp. 267–270). New York, NY, USA: ACM. doi: <http://dx.doi.org/10.1145/2330601.2330666>
- Brooks, C., Greer, J., & Gutwin, C. (2014). The Data-Assisted Approach to Building Intelligent Technology-Enhanced Learning Environments. In J. A. Larusson & B. White (Eds.), *Learning Analytics: From Research to Practice* (pp. 123–156). New York: Springer. doi: http://dx.doi.org/10.1007/978-1-4614-3305-7_7
- Brown, M. (2011). *Learning Analytics: The Coming Third Wave*. Retrieved from <https://net.educause.edu/ir/library/pdf/ELIB1101.pdf>
- Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, 18(6), 683–695. doi: <http://dx.doi.org/10.1080/13562517.2013.827653>
- Davenport, T. H., Harris, J. G., & Morison, R. (2010). *Analytics at work: Smarter decisions, better results*. Boston, MA: Harvard Business Press.
- Driscoll, E., Comm, C. L., & Mathaisel, D. F. X. (2013). A Lesson Plan For Sustainability In Higher Education. *American Journal of Business Education*, 6(2), 255–266. doi: <http://dx.doi.org/10.19030/ajbe.v6i2.7691>
- Ferguson, R., & Shum, S. B. (2011). Learning analytics to identify exploratory dialogue within synchronous text chat. In *1st International Conference on Learning Analytics and Knowledge (LAK '11)* (pp. 99–103). New York, NY, USA: ACM. doi: <http://dx.doi.org/10.1145/2090116.2090130>
- Fritz, J. (2013, April). Using Analytics at UMBC: Encouraging Student Responsibility and Identifying Effective Course Designs. *EDUCAUSE Center for Applied Research*. Retrieved from <https://net.educause.edu/ir/library/pdf/ERB1304.pdf>
- Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), 64–71. doi: <http://dx.doi.org/10.1007/s11528-014-0822-x>
- Gómez-Aguilar, D. A., Hernández-García, Á., García-Peñalvo, F. J., & Therón, R. (2015). Tap into visual analysis of customization of grouping of activities in eLearning. *Computers in Human Behavior*, 47, 60–67. doi: <http://dx.doi.org/10.1016/j.chb.2014.11.001>
- Graydon, B., Urbach-Buholz, B., & Kohen, C. (2011). A study of four textbook distribution models. *Educause Quarterly*, 34(4). Retrieved from <http://www.educause.edu/ero/article/study-four-textbook-distribution-models>
- Haythornthwaite, C., Laat, M. de, & Dawson, S. (2013). Introduction to the Special Issue on Learning Analytics. *American Behavioral Scientist*, 57(10), 1371–1379. doi: <http://dx.doi.org/10.1177/0002764213498850>
- Hilton, J., Gaudet, D., Clark, P., Robinson, J., & Wiley, D. (2013). The Adoption of Open Educational Resources by One Community College Math Department. *The International Review of Research in Open and Distributed Learning*, 14(4), 37–50.
- Hilton, J., Robinson, T. J., Wiley, D., & Ackerman, J. D. (2014). Cost-Savings Achieved in Two Semesters Through the Adoption of Open Educational Resources. *International Review of Research in Open Distributed Learning*, 15(2), 67–84.
- Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative. *Journal of Learning Analytics*, 1(1), 6–47.
- Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender Systems in Technology Enhanced Learning. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender Systems Handbook* (pp. 387–415). Springer US. doi: http://dx.doi.org/10.1007/978-0-387-85820-3_12
- Morris-Babb, M., & Henderson, S. (2012). An experiment in open-access textbook publishing: Changing the world one textbook at a time. *Journal of Scholarly Publishing*, 43(2), 148–155. doi: <http://dx.doi.org/10.3138/jsp.43.2.148>
- Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438–450. doi: <http://dx.doi.org/10.1111/bjet.12152>

Progressing Towards Open Textbooks Learning Analytics System

- Perry, M. (2012). *The college textbook bubble and how the "open educational resources" movement is going up against the textbook cartel*. American Enterprise Institute. Retrieved from <http://www.aei-ideas.org/2012/12/the-college-textbook-bubble-and-how-the-open-educational-resources-movement-is-going-up-against-the-textbook-cartel/>
- Prasad, D., & Usagawa, T. (2014). Towards development of OER derived custom-built open textbooks: A baseline survey of university teachers at the University of the South Pacific. *The International Review of Research in Open and Distributed Learning*, 14(4), 226–247.
- Robinson, T. J., Fischer, L., Wiley, D., & Hilton, J. (2014). The Impact of Open Textbooks on Secondary Science Learning Outcomes. *Educational Researcher*, 43(7), 341–351.
doi: <http://dx.doi.org/10.3102/0013189X14550275>
- Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality Indicators for Learning Analytics. *Educational Technology & Society*, 17(4), 117–132.
- Scheffel, M., Niemann, K., Pardo, A., Leony, D., Friedrich, M., Schmidt, K., ... Kloos, C. (2011). Usage Pattern Recognition in Student Activities. In C. Kloos, D. Gillet, R. Crespo García, F. Wild, & M. Wolpers (Eds.), *Towards Ubiquitous Learning* (Vol. 6964, pp. 341–355). Springer Berlin Heidelberg.
doi: http://dx.doi.org/10.1007/978-3-642-23985-4_27
- Senack, E. (2014). *Fixing the Broken Textbook Market: How Students Respond to High Textbook Costs and Demand Alternatives*. Washington, DC. Retrieved from http://www.washpirg.org/sites/pirg/files/reports/1.27.14_Fixing_Broken_Textbooks_Report.pdf
- Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, 46(5), 30–40.
- Tanes, Z., Arnold, K. E., King, A. S., & Remnet, M. A. (2011). Using Signals for appropriate feedback: Perceptions and practices. *Computers & Education*, 57(4), 2414–2422.
doi: <http://dx.doi.org/10.1016/j.compedu.2011.05.016>
- Wiley, D., Hilton, J., Ellington, S., & Hall, T. (2012). A Preliminary Examination of the Cost Savings and Learning Impacts of Using Open Textbooks in Middle and High School Science Classes. *The International Review of Research in Open and Distributed Learning*, 13(3), 262–276.
- Willis, J. E. (2014). Learning analytics and ethics: A framework beyond utilitarianism. *EDUCAUSE Review Online*. Retrieved from <http://er.educause.edu/articles/2014/8/learning-analytics-and-ethics-a-framework-beyond-utilitarianism>
- Wise, A. F., Zhao, Y., & Hausknecht, S. N. (2014). Learning Analytics for Online Discussions: Embedded and Extracted Approaches. *Journal of Learning Analytics*, 1(2), 48–71.
- Yang, L., & McCall, B. (2014). World education finance policies and higher education access: A statistical analysis of World Development Indicators for 86 countries. *International Journal of Educational Development*, 35, 25–36.
doi: <http://dx.doi.org/10.1016/j.ijedudev.2012.11.002>